

Analytic Theory of Edge Modes in Topological Insulators

Shijun Mao^{1,2}, Yoshio Kuramoto¹, Ken-Ichiro Imura^{1,3}, and Ai Yamakage¹

*Department of Physics, Tohoku University, Sendai 980-8578, Japan*¹

*Department of Physics, Tsinghua University, Beijing 100084, P.R.China*²

*Department of Quantum Matter, AdSM, Hiroshima University, Higashi-Hiroshima 739-8530, Japan*³

(Received August 4, 2010)

Spectrum and wave function of gapless edge modes are derived analytically for a tight-binding model of topological insulators on square lattice. Particular attention is paid to dependence on edge geometries such as the straight (1,0) and zigzag (1,1) edges in the thermodynamic limit. The key technique is to identify operators that combine to annihilate the edge state in the effective one-dimensional (1D) model with momentum along the edge. In the (1,0) edge, the edge mode is present either around the center of 1D Brillouin zone or its boundary, depending on location of the bulk excitation gap. In the (1,1) edge, the edge mode is always present both at the center and near the boundary. Depending on system parameters, however, the mode is absent in the middle of the Brillouin zone. In this case the binding energy of the edge mode near the boundary is extremely small; about 10^{-3} of the overall energy scale. Origin of this minute energy scale is discussed.

KEYWORDS: zigzag lattice, effective Hamiltonian, flat edge spectrum, reentrant edge mode

1. Introduction

Recently, a new class of topological insulator (TI), also referred to as quantum spin Hall (QSH) insulator, has been attracting much interest in both theory¹⁻⁷ and experiments.⁸⁻¹⁰ Different from conventional insulators, QSH insulators have topologically protected helical edge states with the spectrum lying in the bulk insulating gap.^{1,12,13} Two-dimensional (2D) topological insulator has been realized in HgTe/CdTe quantum wells⁸ following theoretical suggestion of Bernevig, Hughes and Zhang⁶ who proposed a useful model, hereafter referred to as the BHZ Model. Helical edge state has been further studied using the resultant continuum model.^{14,15} The BHZ model is not realistic away from the zone center, since it is based on the $\mathbf{k} \cdot \mathbf{p}$ perturbation theory and envelope-function approximation. However, a regularization of the model using the tight-binding scheme has the simple structure for the whole Brillouin zone (BZ), and is suitable for studying the general property of QSH systems. Note that characterization of the topological property of the system requires information of wave functions over the whole BZ.

In this paper, we derive the spectrum of edge modes analytically taking the tight-binding version of the BHZ model for the square lattice. In contrast to previous study of edge modes,^{11,14} we take not only the straight (1,0) edge but also the zigzag (1,1) edge, and make detailed comparison of respective edge modes. This comparison is partially inspired by the remarkable difference between the zigzag and armchair edges in graphene.^{16,17} In contrast to the latter, however, the edges in the present model should not be taken as representing the HgTe/CdTe quantum wells since the short-distance behavior of the model is not realistic. Nevertheless, the intuition gained by the exact solution of the simplified model should provide useful information for understanding more complicated systems.

In a separate paper,¹⁸ we have already derived the spectrum of both straight and zigzag edges relying partially on numerical method. In particular, we have found for the zigzag edge a reentrant mode with tiny binding energy. Since the relevant energy is so small as compared with the overall energy, it is desirable to characterize its nature analytically. Then the origin of tiny binding energy should be clarified. In the present paper, we derive fully analytic expressions of not only the spectrum for both edges, but also momentum regions allowed for the modes. Namely we derive critical momentum where the edge mode merges with bulk excitations. Our analytical method is a systematic generalization of previous ones^{11,19} so that the spectrum can be obtained for general direction of the edge.

This paper is organized as follows: In §2, we review the lattice version of the BHZ model with nearest-neighbor transfer, paying attention to its symmetry property in the 2D BZ. Sections 3 and 4 are devoted to analytic derivation of edge modes for straight and zigzag edges, respectively. We use a systematic method to derive the spectrum in terms of the “annihilator”. Existent regions of edge modes are derived for both straight and zigzag edges. In the zigzag case, we find a novel reentrant edge mode with a tiny binding energy below the bulk spectrum. Finally, we summarize the results and discuss their implication in Sec. 5.

2. Model with particle-hole symmetry

We consider the BHZ model given by the following 4×4 matrix:

$$H(\mathbf{k}) = \begin{bmatrix} h(\mathbf{k}) & 0 \\ 0 & h^*(-\mathbf{k}) \end{bmatrix}, \quad (1)$$

where $\mathbf{k} = (k_x, k_y)$ is a 2D crystal momentum, measured from Γ -point. The lower-right block $h^*(-\mathbf{k})$ is a 2×2 matrix, and is deduced from the upper-left block $h(\mathbf{k})$ by

time reversal transformation. $h(\mathbf{k})$ is parametrized as,

$$h(\mathbf{k}) = \mathbf{d}(\mathbf{k}) \cdot \boldsymbol{\sigma} = \begin{bmatrix} d_z & d_x - id_y \\ d_x + id_y & -d_z \end{bmatrix}, \quad (2)$$

where $d_i(\mathbf{k})$ are given by

$$d_x(\mathbf{k}) = A \sin k_x, \quad d_y(\mathbf{k}) = A \sin k_y, \quad (3)$$

$$d_z(\mathbf{k}) = \Delta - 2B(2 - \cos k_x - \cos k_y), \quad (4)$$

with the lattice constant a set to unity.⁶⁾ We only consider the case $A > 0$, $B > 0$ and $\Delta \geq 0$ in this paper. The same signs of B and Δ are necessary for topological insulator. The solution for $A < 0$ is trivially obtained from that for $H(-\mathbf{k})$. Each row and column of eq. (2) represent spin-orbital states associated with the s -type Γ_6 and the p -type Γ_8 orbitals of the 3D band structure of HgTe and CdTe. Other parameters which appear in Ref.6, i.e., C and D have been set to zero. As a result, the spectrum has the particle-hole symmetry that simplifies the analysis.

The 2×2 -matrix Hamiltonian $h(\mathbf{k})$ is equivalent to the following tight-binding Hamiltonian on the square lattice:

$$h_{\uparrow} = \sum_{I,J} c_{I,J}^{\dagger} \hat{\mathcal{E}}_{I,J} c_{I,J} + \left(c_{I,J}^{\dagger} \hat{t}_x c_{I+1,J} + c_{I,J}^{\dagger} \hat{t}_y c_{I,J+1} + h.c. \right), \quad (5)$$

where we have introduced for each site (I, J) the two-component field:

$$c^{\dagger} = (c_{s\uparrow}^{\dagger}, c_{p\uparrow}^{\dagger}) \quad (6)$$

with orbitals s, p . The tight-binding parameters are given by

$$\hat{\mathcal{E}} = (\Delta - 4B)\sigma_z, \quad (7)$$

$$\hat{t}_x = -i\frac{A}{2}\sigma_x + B\sigma_z, \quad (8)$$

$$\hat{t}_y = -i\frac{A}{2}\sigma_y + B\sigma_z, \quad (9)$$

each of which is a 2×2 matrix. The down spin part h_{\downarrow} corresponding to the lower-right block of eq.(1) is obtained by replacing \hat{t}_i with its complex conjugation \hat{t}_i^* .

We consider the BHZ model over the whole BZ of the square lattice. Bulk energy $E_b(\mathbf{k})$ is written as

$$E_b(\mathbf{k}) = \pm \left\{ A^2 (\sin^2 k_x + \sin^2 k_y) + [\Delta - 4B + 2B (\cos k_x + \cos k_y)]^2 \right\}^{1/2}, \quad (10)$$

which is symmetric with respect to positive (conduction) and negative (valence) energy bands, being the signature of the particle-hole symmetry. As one varies mass parameter Δ , there appear four gap closing points, namely at $\Gamma = (0, 0)$, $X = (\pi, 0)$, $X' = (0, \pi)$ and $M = (\pi, \pi)$. Note that these points are invariant against time-reversal operation. Gap closing at Γ occurs when $\Delta = 0$, whereas the gap closing at X and X' occurs simultaneously when $\Delta = 4B$, at M when $\Delta = 8B$.

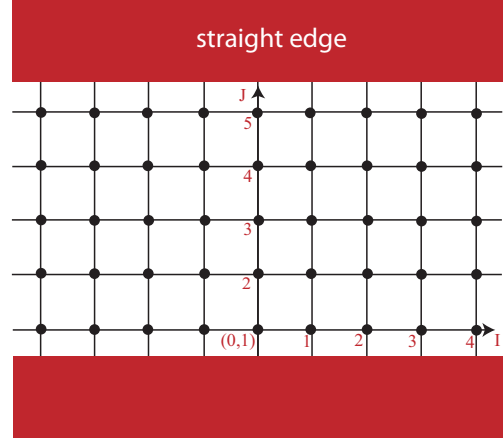


Fig. 1. Straight edge lattice ribbon with two boundaries in $(1,0)$ direction.

3. Straight edge

3.1 Effective one-dimensional form

Let us first consider the geometry (see Fig. 1), in which electrons are confined to N_r rows in a strip between $y = 1$ and $y = N_r$, i.e., the two edges are along x -axis. Translational invariance along x -axis allows for constructing a 1D Bloch state with a crystal momentum $k = k_x$:

$$|k, J\rangle = \frac{1}{\sqrt{N_c}} \sum_I e^{ikI} |I, J\rangle, \quad (11)$$

where I, J represents lattice site, N_c is the number of sites along the x -axis, and $|I, J\rangle = c_{I,J}^{\dagger} |0\rangle$. In order to introduce the edges, it is convenient to rewrite eq. (5) in form of a hopping Hamiltonian between neighboring rows. In terms of the two-component creation and annihilation operators $c_J^{\dagger}(k)$ and $c_J(k)$ associated with Bloch state (11), one can rewrite eq. (5) as

$$h_{\uparrow} = \sum_k h_{01}(k), \quad (12)$$

$$h_{01}(k) = \sum_J c_J^{\dagger}(k) \hat{\mathcal{E}}(k) c_J(k) + \sum_J \left[c_J^{\dagger}(k) \hat{t}_y c_{J+1}(k) + h.c. \right] \quad (13)$$

where $\hat{\mathcal{E}}(k)$ is given by

$$\hat{\mathcal{E}}(k) = A \sin k \sigma_x + (\Delta_B + 2 \cos k) \sigma_z, \quad (14)$$

Here and in the following, we use the notation $\Delta_B \equiv \Delta - 4B$, and take the energy unit so that $B = 1$. The corresponding Schrödinger equation is given by

$$\hat{\mathcal{E}}(k) \Psi_J + \hat{t}_y \Psi_{J-1} + \hat{t}_y^{\dagger} \Psi_{J+1} = E_{\uparrow}(k) \Psi_J \quad (15)$$

where Ψ_J is the two-component amplitude with row index J . The straight edges along the $J = 1$ row and $J = N_r$ row can be implemented by open boundary condition $\Psi_0 = \Psi_{N_r+1} = 0$.

It has been shown¹⁴⁾ in the continuum approximation of eq.(15) that there appear modes, which are localized on both edges of the system, with the particle-hole symmetric spectrum and a minimum gap at $k = 0$. The origin

of the gap is the overlap of wave functions on different edges. As N_r tends to infinity, the gap disappears since the overlap becomes zero. In this limit one can identify edge modes localized on either edge. In the thermodynamic limit, the spectrum $E(k)$ of an edge mode near $J = 1$ becomes an odd function of k , and the other edge mode has the spectrum $-E(k)$. The continuum approximation of Ref.14 has a limited validity, but emergence of the zero mode at $k = 0$ in the thermodynamic limit $N_r \rightarrow \infty$ is guaranteed by the time-reversal and particle-hole symmetries.^{11,19} In the following, we deal with the thermodynamic limit.

3.2 Separation into Hermitian and annihilating parts

Our strategy to obtain the edge spectrum is best illustrated in the straight edge. Although the spectrum in this case has already been obtained in the literature,^{11,19} we here present our way of derivation that can be extended to the zigzag edge. We try an edge state solution with property: $\Psi_{J+1} = \lambda \Psi_J \equiv \lambda^{J+1} \Psi$ with $|\lambda| < 1$,^{19,20} and derive the vector Ψ . Then eq.(15) can be written in the following form:

$$\left[\hat{\mathcal{E}}(k) + \lambda \hat{t}_y^\dagger + \frac{1}{\lambda} \hat{t}_y \right] \Psi \equiv P_{01}(\lambda, k) \Psi = E_\uparrow(k) \Psi, \quad (16)$$

where $P_{01}(\lambda, k)$ can be rearranged as

$$P_{01}(\lambda, k) = \gamma \cdot \sigma, \quad (17)$$

with components

$$\gamma_x = A \sin k \quad (18)$$

$$\gamma_y = i \frac{A}{2} (\lambda - \lambda^{-1}) \quad (19)$$

$$\gamma_z = (\lambda + \lambda^{-1}) + \Delta_B + 2 \cos k \quad (20)$$

For a general complex vector γ , the eigenvalues of P_{01} are also complex. However, we obtain real energy $E_\uparrow(k)$ in eq.(16) if one of the following conditions is met:

- (a) all components of γ are real (including zero);
- (b) the nonzero complex components combine to give zero when acting on the edge state. Such combination of operators is referred to as annihilator.

The condition (a) is not relevant here, because we would then have two eigenvalues $\pm E_\uparrow(k)$ and corresponding two eigenfunctions for a given k . Actually we should have only one edge mode for $h_\uparrow(k)$. Hence we have to accept the condition (b), and separate P_{01} into a Hermitian part that gives the spectrum by diagonalization, and the rest that makes up the annihilator.

From eq.(19) and (20), coefficients γ_y and γ_z are both real only if $|\lambda| = 1$. On the other hand, due to time-reversal symmetry and particle-hole symmetry of the system, the eigenvalue $E_\uparrow(k)$ is an odd function of k in the thermodynamic limit.^{11,14} Therefore, σ_y and σ_z must both belong to the annihilator, and the only component to be diagonalized is σ_x . The annihilator corresponds to either $\sigma_y + i\sigma_z$ or $\sigma_y - i\sigma_z$, depending on the eigenvalue ± 1 of σ_x .

Accordingly, we decompose $P_{01}(\lambda, k)$ as

$$P_{01}(\lambda, k) = H_{01} + F_{01}, \quad (21)$$

$$H_{01} = A \sin k \sigma_x, \quad (22)$$

$$F_{01} = \gamma_y \sigma_y + \gamma_z \sigma_z, \quad (23)$$

where F_{01} should form the annihilator. Namely, we impose the relation $\gamma_z = si\gamma_y$, according to the eigenvalue $s = \pm 1$ of σ_x . The relation is equivalent to

$$\left(1 + \frac{sA}{2}\right)\lambda + \left(1 - \frac{sA}{2}\right)\lambda^{-1} + m_k = 0, \quad (24)$$

where we have introduced the notation:

$$m_k = \Delta_B + 2 \cos k. \quad (25)$$

By solving eq.(24) with $A \neq 2$, we obtain

$$\lambda_\pm(s) = \frac{-m_k \pm \sqrt{m_k^2 + A^2 - 4}}{2 + sA}, \quad (26)$$

with the relation

$$\lambda_\pm(-s) = 1/\lambda_\mp(s). \quad (27)$$

Provided the eigenvalue equation

$$H_{01} \Psi = E_\uparrow(k) \Psi. \quad (28)$$

is satisfied, then equality $F_{01} \Psi = 0$ follows with eq.(24). In this way the edge spectrum is simply derived as

$$E_\uparrow(k) = sA \sin k. \quad (29)$$

with the eigenstate written as

$$\Psi(s) = \begin{pmatrix} 1 \\ s \end{pmatrix}. \quad (30)$$

Note that only one of $s = \pm 1$ is relevant, as discussed below.

The boundary condition $\Psi_{J=0} = 0$ requires the edge state to have the form

$$\Psi_J(s) = [\lambda_+(s)^J - \lambda_-(s)^J] \Psi(s), \quad (31)$$

apart from the normalization factor. Since the wave function should decay as J increases, the edge state can only be realized with $|\lambda_\pm(s)| < 1$. Because of the relation eq.(27), there is at most one s for given k that describes the edge mode with the spectrum eq.(29). Hence there is only one edge state per spin and per momentum.

In a similar manner, we obtain another edge mode for h_\downarrow . The corresponding energy $E_\downarrow(k)$ is given by

$$E_\downarrow(k) = E_\uparrow(-k) = -E_\uparrow(k), \quad (32)$$

where the first equality corresponds to the time-reversal symmetry. The particle-hole symmetry connecting the rightmost and leftmost sides with the same k involves different spins. Note that the Kramers pair has the same $\lambda_\pm(s)$ for given k and $-k$. In certain range of k , however, the edge modes do not exist. This problem is studied in detail in the next section.

3.3 Allowed momentum range for edge modes

A pair of gapless edge modes per edge are always present in the topological insulator phase (TI) that appears for $0 < \Delta < 8$, i.e., $|\Delta_B| < 4$ in the BHZ model.

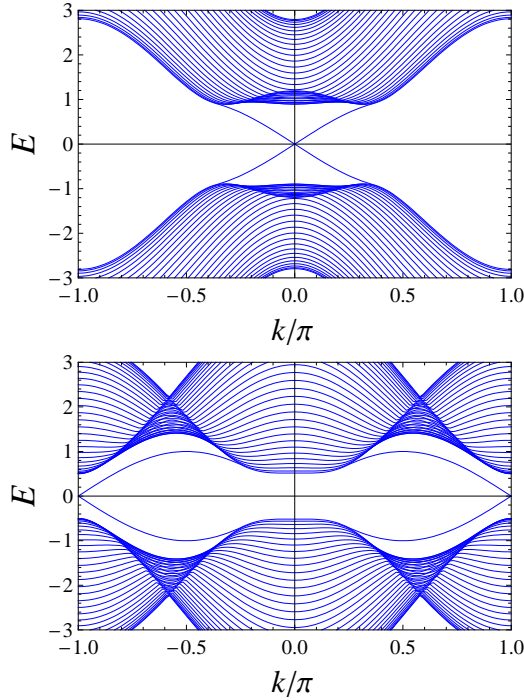


Fig. 2. 1D energy bands in the straight edge: TI-1 (top) with $\Delta = 1.2$ and TI-2 (bottom) with $\Delta = 4.5$. In both cases, we set $A = 1$. The edge modes have the spectrum $\pm A \sin k$ in the existent range of k .

However, the number of zero points of helical edge states should be an odd multiple of two, including the degeneracy, in the 1D BZ.¹²⁾ If the edge modes with the spectrum $E = \pm A \sin k$ were present for the whole BZ, the zero points amount to four (an even multiple of two) which violates the topological stability.²¹⁾ Hence the edge modes must merge into bulk excitations at finite κ_m , and degenerate zero points occur either at $k = 0$ or $k = \pi$, but not at both. Note that $k = 0$ and $k = \pi$ are two time reversal invariant momenta in 1D BZ.

With $\Delta_B = 0$, the energy gap closes at X points $(\pi, 0)$ and $(0, \pi)$ in the 2D BZ, as seen from eq.(10). Let us classify the case of $-4 < \Delta_B < 0$ as TI-1, and the case of $0 < \Delta_B < 4$ as TI-2. Figure 2 shows the spectrum of edge modes in the TI-1 (upper panel) and TI-2 (lower panel). In the TI-1, the edge modes intersect at $k = 0$, whereas in TI-2 they meet at $k = \pi$. Also shown is approximate bulk spectrum in the system that is derived from $E_b(k, k_y)$ by fixing k_y to $2\pi n/N_r$ with $n = 1, \dots, N_r$. Each curve for the bulk spectrum corresponds to integer n .

We now derive the pair $\pm k_m$ of momentum where the edge mode merges with bulk excitations. In the following, we always assume $k \geq 0$ for simplicity. The merging occurs when the larger of $|\lambda_{\pm}(s)|$ becomes unity. Note that such $\lambda_{\pm}(s)$ is real, since otherwise $|\lambda_{\pm}(s)|$ is independent of m_k , and hence of k . For complex $\lambda_{\pm}(s)$ we obtain from eq.(26),

$$|\lambda_{\pm}(s)|^2 = \frac{2 - sA}{2 + sA}, \quad (33)$$

which is less than unity only with $s = 1$. In this section we deal with the case $0 < A < 2$ that allows complex

$\lambda_{\pm}(s)$. Then, edge modes are assured to be present for such k with complex $\lambda_{\pm}(1)$, which is simply written as λ_{\pm} hereafter. Edge modes in the case of $A > 2$ will be discussed in §5.

TI-1 regime

In this case gapless points are present at $k = 0$. Let us assume $m_k = \Delta_B + 2 \cos k < 0$ when merging occurs at $k = k_{m1}$. Then we obtain $|\lambda_+| > |\lambda_-|$ for real λ_{\pm} , and the condition for merging is reduced to

$$\lambda_+ = \frac{-m_k + \sqrt{m_k^2 + A^2 - 4}}{2 + A} = 1. \quad (34)$$

from eq.(26). Then we obtain

$$\cos k_{m1} = 1 - \Delta/2, \quad (35)$$

which justifies the assumption $m_k < 0$. One can check that there is no solution if we assume $m_k > 0$. We note that eq.(35) has already been obtained by König et al.¹¹⁾

The condition for merging is also to have the same energy as the lowest bulk excitation for given k . The minimum of $E_b(k, k_y)$ may occur either at $k_y = 0$ or $k_y = \pm \arccos[2m_k/(A^2 - 4)]$ depending on the value of k . If the threshold of bulk excitations with $k = k_{m1}$ occurs at $k_y = 0$, merging momentum k_{m1} is simply obtained from eq.(10) as

$$\Delta_B + 2 + 2 \cos k = m_k + 2 = 0, \quad (36)$$

which is consistent with eq.(35). Namely, the condition $\lambda_+ = 1$ is equivalent to having the same energy for edge mode and for the minimum of bulk excitations. Furthermore, it is easily seen that the group velocity $A \cos k_{m1}$ at merging point is common to both edge mode and the lowest bulk excitation. Namely, the edge mode vanishes at such k that it has the common tangent with the threshold of bulk excitations $E_b(k, k_y)$ with $k_y = 0$.

The crossings of 1D energy bands in Fig.2 indicates the transition from the lowest bulk excitation $E_b(k, k_y)$ with $k_y = 0$ to with $k_y = \pm \arccos[2m_k/(A^2 - 4)]$. The critical momentum k_{c1} satisfies the condition

$$\partial^2 E_b(k_{c1}, k_y) / \partial k_y^2|_{k_y=0} = 0, \quad (37)$$

which gives the solution

$$\cos k_{c1} = 1 - \frac{\Delta}{2} + \frac{A^2}{4}. \quad (38)$$

The threshold has $k_y = 0$ for $k > k_{c1} > 0$. By comparing with eq.(35), we find $\cos k_{c1} > \cos k_{m1}$, which means $0 \leq k_{c1} < k_{m1}$. Namely, merging with bulk excitations indeed occurs in the range where $k_y = 0$ corresponds to the threshold.

TI-2 regime

In this case edge modes exist around $k = \pi$. Let us assume $m_k > 0$ when merging occurs at $k = k_{m2} > 0$. The condition for merging is now given by

$$\lambda_- = \frac{-m_k - \sqrt{m_k^2 + A^2 - 4}}{2 + A} = -1, \quad (39)$$

which gives $m_k = 2$ as the solution, or

$$\cos k_{m2} = 1 - \Delta_B/2 = 3 - \Delta/2. \quad (40)$$

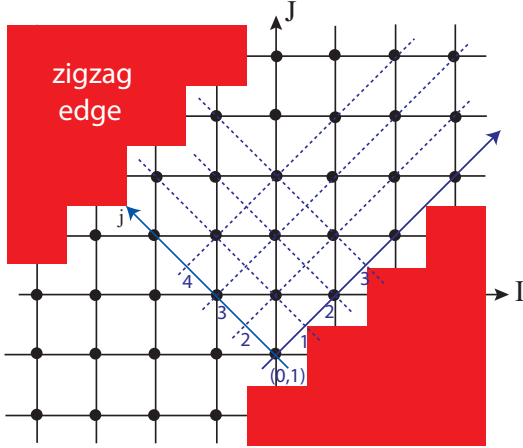


Fig. 3. Zigzag edge lattice ribbon, with two boundaries in (1,1) direction.

In the TI-2 regime, the threshold of bulk excitations occurs at $k_y = \pi$ or $k_y = \pm \arccos[2m_k/(A^2 - 4)]$. It can be checked that the bulk energy at $\mathbf{k} = (k_{m2}, \pi)$ becomes the same as the edge mode with the condition (40). Hence merging occurs with bulk excitations $E_b(k, k_y)$ with $k_y = \pi$.

The critical momentum $k = k_{c2}$, below which the minimum of $E_b(k, k_y)$ no longer occurs at $k_y = \pi$, can be obtained by condition

$$\partial^2 E_b(k_{c2}, k_y) / \partial k_y^2|_{k_y=\pi} = 0, \quad (41)$$

which gives the solution

$$\cos k_{c2} = 1 - \frac{\Delta_B}{2} - \frac{A^2}{4}. \quad (42)$$

Thus we have the relation $\cos k_{c2} < \cos k_{2m}$, or $k_{c2} > k_{2m}$. Hence merging indeed occurs in the range where $k_y = \pi$ corresponds to the threshold. In this way, we have quantified important characteristics of the edge modes shown in Fig.2.

4. Zigzag edge

4.1 Effective one-dimensional form

Let us now consider a zigzag edge geometry, as illustrated in Fig.3. Electrons in a zigzag edge geometry are confined in a diagonal strip: $1 \leq y - x \leq N_r$, provided the edges are placed at $y - x = 1$ and $y - x = N_r$, normal to the $(1, -1)$ -direction. The translational invariance remains along the $(1, 1)$ -direction, where the conserved momentum is given by $p = (k_x + k_y)/\sqrt{2}$. For notational convenience we introduce $\kappa = p/\sqrt{2}$ with $-\pi/2 < \kappa \leq \pi/2$ and define the new basis set as

$$|\kappa, j\rangle = \frac{1}{\sqrt{N_c}} \sum_I \exp[i\kappa(2I + j)] |I, I + j\rangle, \quad (43)$$

where the site summation goes along the $(1, 1)$ direction. The phase factor is so chosen that it becomes unity for the state $|-J, J\rangle$. Then, the amplitude Φ_j in this basis

satisfies the Schrödinger equation analogous to eq.(15):

$$\hat{\mathcal{E}}\Phi_j + \hat{t}_{11}(\kappa)\Phi_{j-1} + \hat{t}_{11}^\dagger(\kappa)\Phi_{j+1} = E_\uparrow(\kappa)\Phi_j \quad (44)$$

where $\hat{\mathcal{E}}$ has been defined by eq.(7) and hopping matrix $\hat{t}_{11}(\kappa)$ is given by

$$\hat{t}_{11}(\kappa) = \frac{i}{2}Ae^{-i\kappa}\sigma_x - \frac{i}{2}Ae^{i\kappa}\sigma_y + 2\cos\kappa\sigma_z \quad (45)$$

We impose the boundary condition: $\Phi_0 = \Phi_{N_r+1} = 0$, which is consistent with the zigzag edge geometry. Assuming eigenstate of eq.(44) with property $\Phi_j = \lambda\Phi_{j-1} = \lambda^j\Phi$, where $|\lambda| < 1$,^{19, 20)} we obtain

$$\left(\hat{\mathcal{E}} + \lambda\hat{t}_{11}^\dagger(\kappa) + \lambda^{-1}\hat{t}_{11}(\kappa)\right)\Phi = P_{11}(\lambda, \kappa)\Phi = E_\uparrow(\kappa)\Phi \quad (46)$$

For later reference purpose, we write the bulk energy E_b in terms of variables $\kappa = (k_x + k_y)/2$ and $\xi = (k_x - k_y)/2$. From eq.(10) we obtain

$$E_b(\kappa, \xi) = \pm \left[2A^2 (\sin^2 \kappa \cos^2 \xi + \cos^2 \kappa \sin^2 \xi) + (\Delta_B + 4 \cos \kappa \cos \xi)^2 \right]^{1/2}. \quad (47)$$

4.2 Derivation of spectrum in thermodynamic limit

We will separate P_{11} into the Hermitian part H_{11} and the corresponding annihilator F_{11} . The separation now is not straightforward in contrast with the case of straight edge. As a preliminary, we introduce the following matrices:

$$\sigma_X = (\sigma_x + \sigma_y) / \sqrt{2}, \quad (48)$$

$$\sigma_Y = (\sigma_y - \sigma_x) / \sqrt{2}. \quad (49)$$

Then we obtain

$$\hat{t}_{11}(\kappa) = \frac{A}{\sqrt{2}} (\sin \kappa \sigma_X - i \cos \kappa \sigma_Y) + 2 \cos \kappa \sigma_z. \quad (50)$$

We note that the spectrum of each edge mode is an odd function of κ . Then we introduce a variable θ , which is an odd function of κ , and make the following transformation:

$$\sigma_{\theta x} = \cos \theta \sigma_X + \sin \theta \sigma_z, \quad (51)$$

$$\sigma_{\theta z} = \cos \theta \sigma_z - \sin \theta \sigma_X, \quad (52)$$

and rewrite as $\sigma_{\theta y} = \sigma_Y$. They keep the commutation property:

$$[\sigma_{\theta x}, \sigma_{\theta y}] = 2i\sigma_{\theta z}, \quad (53)$$

and analogous cyclic ones that are the same as the original Pauli matrices. Then we obtain

$$\hat{\mathcal{E}} = \Delta_B (\sin \theta \sigma_{\theta x} + \cos \theta \sigma_{\theta z}), \quad (54)$$

$$\lambda\hat{t}_{11}^\dagger + \lambda^{-1}\hat{t}_{11} = \gamma_{\theta x}\sigma_{\theta x} + \gamma_{\theta y}\sigma_{\theta y} + \gamma_{\theta z}\sigma_{\theta z}, \quad (55)$$

where

$$\gamma_{\theta x} = (\lambda + \lambda^{-1}) \left(\frac{A}{\sqrt{2}} \sin \kappa \cos \theta + 2 \cos \kappa \sin \theta \right), \quad (56)$$

$$\gamma_{\theta y} = i(\lambda - \lambda^{-1}) \frac{A}{\sqrt{2}} \cos \kappa, \quad (57)$$

$$\gamma_{\theta z} = (\lambda + \lambda^{-1}) \left(2 \cos \kappa \cos \theta - \frac{A}{\sqrt{2}} \sin \kappa \sin \theta \right). \quad (58)$$

We choose Φ as eigenstate of $\sigma_{\theta x}$ since the coefficient $\Delta_B \sin \theta$ in eq.(54) is an odd function of κ . Then terms with $\sigma_{\theta y}$ and $\sigma_{\theta z}$ must combine to form the annihilator. Furthermore, since the coefficient of $\sigma_{\theta x}$ must be real, and be an odd function of κ , we require $\gamma_{\theta x} = 0$. This condition determines θ in terms of κ as

$$\sin \theta = -\frac{\tan \kappa}{\sqrt{\tan^2 \kappa + 8/A^2}}, \quad (59)$$

$$\cos \theta = \frac{\sqrt{8}/A}{\sqrt{\tan^2 \kappa + 8/A^2}}. \quad (60)$$

In this way, we decompose $P_{11} = H_{11} + F_{11}$ in eq.(46) as follows:

$$H_{11} = \Delta_B \sin \theta \sigma_{\theta x}, \quad (61)$$

$$F_{11} = (\Delta_B \cos \theta + \gamma_{\theta z}) \sigma_{\theta z} + \gamma_{\theta y} \sigma_{\theta y}. \quad (62)$$

By diagonalizing H_{11} , the eigenenergy $E_{\uparrow}(\kappa)$ is derived as

$$E_{\uparrow}(\kappa) = s \Delta_B \sin \theta, \quad (s = \pm 1), \quad (63)$$

where only one of the signs \pm is relevant, as derived shortly. Note that the spectrum has a form analogous to the case of straight edge given by eq.(29). The condition for F_{11} to form the annihilator is given by

$$\Delta_B \cos \theta + \gamma_{\theta z} = i s \gamma_{\theta y} \quad (64)$$

which determines λ for the edge mode as

$$\lambda_{\pm}(s) = \frac{1}{2 \cos \kappa} \cdot \frac{-\Delta_B \cos \theta \pm \sqrt{R}}{2 / \cos \theta + s A / \sqrt{2}}, \quad (65)$$

with

$$R = \Delta_B^2 \cos^2 \theta - 2 \cos^2 \kappa (8 \cos^{-2} \theta - A^2). \quad (66)$$

In the case of $R < 0$, we obtain complex $\lambda_{\pm}(s)$ with absolute value

$$|\lambda_{\pm}(s)|^2 = \frac{\sqrt{8} - s A \cos \theta}{\sqrt{8} + s A \cos \theta}, \quad (67)$$

which is less than unity only for $s = 1$, and positive for $A^2 \leq 8$. Therefore, the edge mode must have $s = 1$ in eq.(63), and the group velocity is positive (negative) in TP-1 (TP-2) regime. A special case occurs with $\theta = \pm \pi/2$ that corresponds to $\kappa = \mp \pi/2$ according to eqs.(59) and (60). Actually eq.(65) gives $\lambda_{\pm} = \pm i$ at $\kappa = \pi/2$. Thus, at the boundary of the 1D BZ, there are no edge modes since $|\lambda_{\pm}| = 1$. The neighborhood of this special point has $|\lambda_{\pm}| < 1$, and there should be an edge mode. We emphasize that this property is independent of Δ and A , and is specific to the zigzag edge.

Due to the time-reversal symmetry, we obtain the Kramers partner from h_{\downarrow} with the spectrum $E_{\downarrow}(\kappa) = -\Delta_B \sin \theta$. Figure 4 shows the edge modes together with bulk excitations for the TI-1 and TI-2, and the boundary case $\Delta_B = 0$. The bulk spectrum illustrated is obtained from $E_b(\kappa, \xi)$ as a function of κ with fixed ξ . In both regimes, the edge mode becomes gapless at $\kappa = 0$.

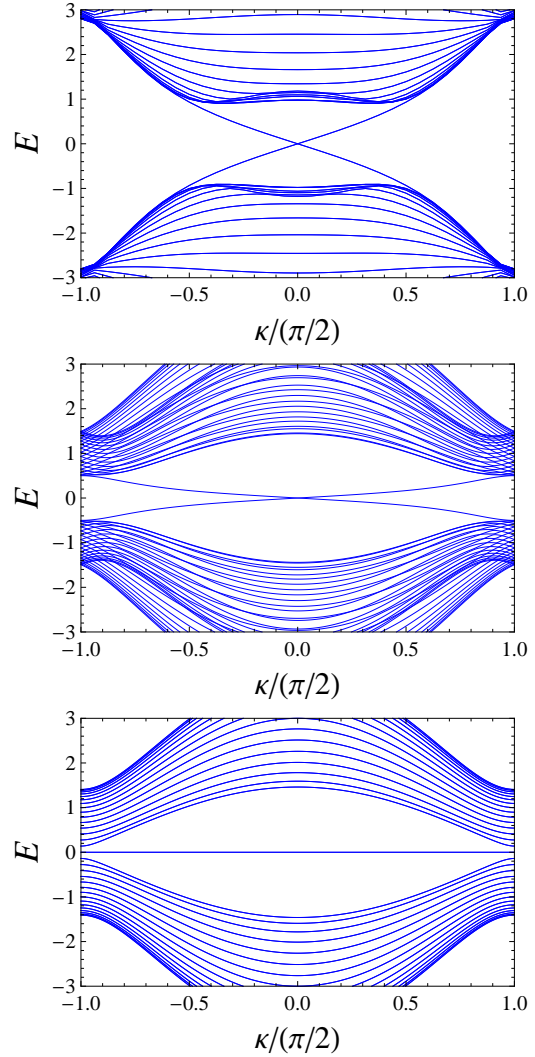


Fig. 4. Spectrum of zigzag edge modes. The top panel is for TI-1 state with $\Delta = 1.2$, the center is for TI-2 state with $\Delta = 4.5$. The bottom panel shows the boundary case $\Delta = 4$. All cases have $A = 1$.

Hence only $\kappa = 0$ is the relevant point where a pair of edge modes are degenerate by time-reversal invariance. We note that the spectrum becomes completely flat with $\Delta_B = 0$.¹⁸⁾

In the low momentum region, the spectrum tends to the linear dispersion

$$E_{\uparrow, \downarrow}(\kappa) = \mp \frac{1}{\sqrt{8}} A \Delta_B \kappa = \mp \frac{1}{4} A \Delta_B p. \quad (68)$$

Especially, with $\Delta_B = \pm 4$, the spectrum becomes the same as the corresponding modes in the straight edge. Thus we find that the system acquires the axial symmetry only in the case of $\Delta_B = \pm 4$ even in the long-wavelength limit. This is not surprising since the difference in the edge shape remains even for long wavelength.

4.3 Allowed momentum range for edge modes

Let us first consider the edge mode in TI-1 regime corresponding to right-going edge mode (see Figure 2 top panel), and $\lambda_{\pm} = \lambda_{\pm}(1)$. We restrict to the region of

positive κ . Since complex λ always has a pair of solutions with $|\lambda_{\pm}| \leq 1$, the merging momentum $\kappa_m < \pi/2$ can be obtained from the condition

$$\lambda_+ = 1, \quad (69)$$

which is equivalent to

$$(8 - A^2) \cos^2 \kappa_m + 2\Delta_B \cos \kappa_m + A^2 = 0, \quad (70)$$

according to eq.(65). Note that there is no solution for $0 \leq \cos \kappa_m \leq 1$ in the case of $A^2 \geq 8$. The relevant solution in the case of $A^2 < 8$ is given by

$$\cos \kappa_m = \frac{-\Delta_B \pm \sqrt{\Delta_B^2 - A^2(8 - A^2)}}{8 - A^2}, \quad (71)$$

The critical value of Δ_B beyond which no real κ_m exists is given by

$$\Delta_{4c} = -A\sqrt{8 - A^2} < 0, \quad (72)$$

where we consider only the case $0 < A < 2$ as in the the straight edge. According to eq.(72), we have three cases:

- (i) no solution for $\kappa_m < \pi/2$ with $\Delta_B > \Delta_{4c}$;
- (ii) single κ_m with $\Delta_B = \Delta_{4c}$;
- (iii) two solutions $\kappa_{m1} < \kappa_{m2}$ with $\Delta_B < \Delta_{4c}$.

Edge modes in the case of $A > 2$ will be discussed in §5.

The threshold of bulk excitations can occur either at $\xi = 0$ or $\xi \neq 0$ depending on κ . The critical value κ_c separating the two cases is determined by the condition

$$\partial^2 E_b(\kappa_c, \xi) / \partial \xi^2|_{\xi=0} = 0, \quad (73)$$

which can be reduced to

$$2(A^2 - 4) \cos^2 \kappa_c - 2\Delta_B \cos \kappa_c - A^2 = 0. \quad (74)$$

Then we obtain

$$\cos \kappa_c = \frac{\Delta_B \pm \sqrt{\Delta_B^2 + 2A^2(A^2 - 4)}}{2(A^2 - 4)}. \quad (75)$$

In the case of $\xi \neq 0$, the momentum ξ at the threshold satisfy the condition

$$\cos \xi = \frac{-2\Delta_B \cos \kappa}{4 + (4 - A^2) \cos 2\kappa}. \quad (76)$$

At the zone boundary $\kappa = \pi/2$, eq.(76) gives $\xi = \pm\pi/2$ as the solution. Here the edge mode has the energy $E_{\uparrow}(\pi/2) = -\Delta_B$, and the lowest bulk excitation has the same energy $E_b(\pi/2, \pm\pi/2) = -\Delta_B$. Namely, $\kappa = \pi/2$ is always a merging point in the zigzag edge for any parameter settings.

In the following, we analyze the spectrum of the edge mode near the merging momentum according to classification (i), (ii), (iii) given above.

Edge modes for the whole BZ

Let us first consider the case (i): $\Delta_B > \Delta_{4c}$. With $A = 1$, we obtain $\Delta_{4c} = -\sqrt{7}$, i.e., $\Delta_c \sim 1.354$ from eq.(72). Figure 5 shows $|\lambda_{\pm}|$ and the energy of the edge mode relative to the threshold of bulk excitations in this case with $\Delta = 1.4$.

At the zone boundary, the edge mode merges with bulk

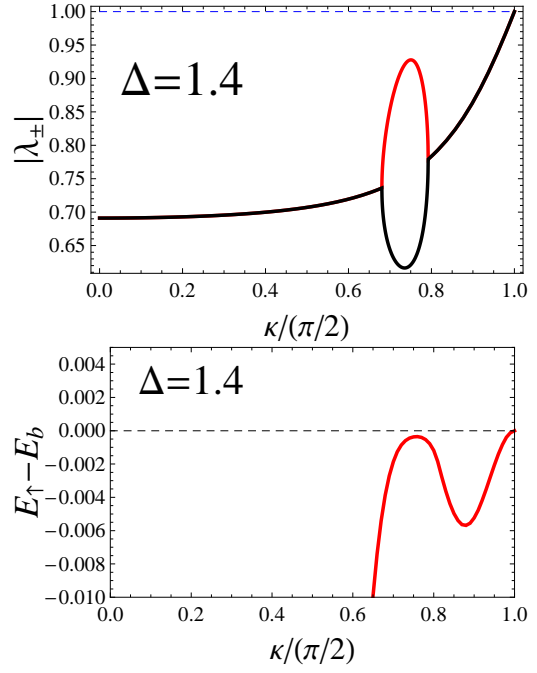


Fig. 5. The parameter $|\lambda_{\pm}|$ and energy difference $E_{\uparrow} - E_b$ between edge mode and the threshold of bulk excitations as a function of κ for $\Delta = 1.4$ with $A = 1$. For complex λ , we obtain $|\lambda_+| = |\lambda_-|$. In the region of κ with real λ_{\pm} , $E_{\uparrow} - E_b$ becomes almost zero, but is marginally negative.

excitations. The difference of energies is expanded as

$$E_{\uparrow}(\kappa) - E_b(\kappa, \frac{\pi}{2}) = \frac{A^2}{\Delta_B} (\kappa - \frac{\pi}{2})^2 + O\left((\kappa - \frac{\pi}{2})^4\right). \quad (77)$$

Note that only even order terms appear in the expansion since both the lowest bulk excitation E_b and edge mode energy E_{\uparrow} are symmetric around $\pi/2$.

Edge modes with critical momentum κ_m

Next we consider the critical case (ii) characterized by single κ_m with $\Delta_B = \Delta_{4c}$. Figure 6 shows $|\lambda_{\pm}|$ and the energy difference. By comparing eqs.(71) and (75), we obtain

$$0 < \kappa_{c1} < \kappa_m < \kappa_{c2} < \pi/2. \quad (78)$$

Hence we find $\xi = 0$ for the bulk momentum at the merging point.

Edge modes with reentrance

We finally consider the case (iii): two solutions $\kappa_{m1} < \kappa_{m2}$ with $\Delta_B < \Delta_{4c}$. From eqs.(71) and (75) we obtain the relation

$$0 < \kappa_{c1} < \kappa_{m1} < \kappa_{m2} < \kappa_{c2} < \pi/2 \quad (79)$$

Hence we obtain the bulk momentum $\xi = 0$ for both merging points. Then we expand $E_{\uparrow} - E_b$ around merging points κ_{mi} ($i = 1, 2$) as follows:

$$E_{\uparrow}(\kappa) - E_b(\kappa, 0) = a_i(\kappa - \kappa_{mi})^2 + O((\kappa - \kappa_{mi})^3), \quad (80)$$

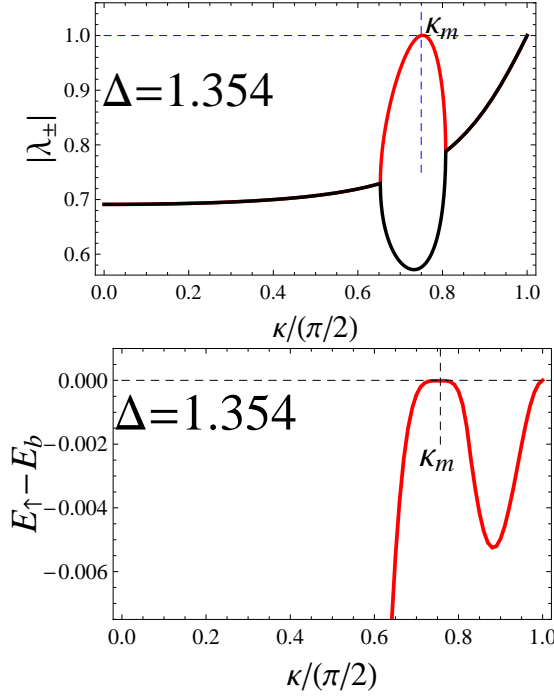


Fig. 6. The same quantities as in Fig.5 but with $\Delta = \Delta_c = 4 - \sqrt{7} \sim 1.354$. Note the scale of the ordinate in the lower panel, showing the minute energy difference as compared with the overall energy scale of the system.

where the expansion coefficient is calculated as

$$a_i = \frac{4(A^4 - 8A^2 + \Delta_B^2)}{A^2 \Delta_B} \times \left[\frac{A^2 - 4}{A^2 - 8} + (-1)^i \frac{4\sqrt{A^4 - 8A^2 + \Delta_B^2}}{\Delta_4(A^2 - 8)} \right]^{1/2} \quad (81)$$

Hence we have proven that the threshold of bulk excitation shares the same energy and velocity with edge mode, since the lowest order term of expansion is of second order.

At critical value of $\Delta_B = \Delta_{4c}$, we obtain $\kappa_m = \kappa_{m1} = \kappa_{m2}$, and both second and third order terms in eq.(80) tend to zero. Then the expansion around κ_m begins from fourth order. This explains the nearly flat shape of $E_\uparrow - E_b$ in Fig.6.

Presence of two merging points causes novel phenomenon in zigzag edge mode. Figure 7 shows the parameter $|\lambda_\pm|$ and the energy difference as in previous cases. In the region $0 < \kappa < \kappa_{m1}$ and $\kappa_{m2} < \kappa < \frac{\pi}{2}$, we obtain $|\lambda_\pm| < 1$ and $E_\uparrow - E_b < 0$. Namely, the edge mode has two separate momentum ranges of its existence. The normal edge state (I) starts from zone center ($\kappa = 0$), and vanishes at an intermediate point $\kappa = \kappa_{m1}$. In addition the reentrant part (II) appears near the zone boundary: $\kappa_{m2} < \kappa < \frac{\pi}{2}$. The reentrant edge mode has only marginally lower energy than lowest bulk excitation, which approximatively obeys $E_\uparrow - E_b \propto (\kappa - \kappa_{m2})^2$. In the momentum range $\kappa_{m1} \leq \kappa \leq \kappa_{m2}$, the edge mode disappears since one of $|\lambda_\pm|$ exceeds unity.

Let us summarize the results for different Δ with fixed A shown in Figs. 5, 6 and 7. As Δ increases from zero,

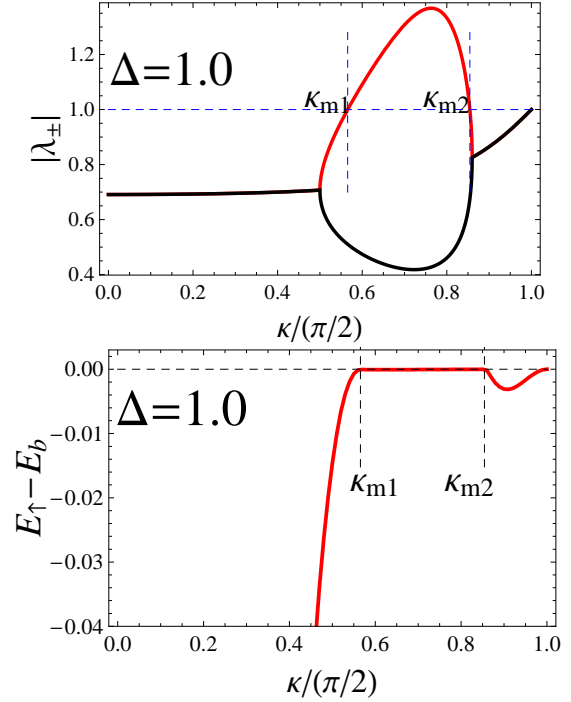


Fig. 7. The same quantities as in Fig.5 but with $\Delta = 1$. There is no edge mode for κ between κ_{m1} and κ_{m2} where one of $|\lambda_\pm|$ exceeds unity.

the two separate regions for the edge modes widen in momentum space simultaneously. Namely, both κ_{m1} and $\pi/2 - \kappa_{m2}$ increase, while $\kappa_{m2} - \kappa_{m1}$ decreases with increasing Δ . At critical Δ_c , the two regions merge ($\kappa_{m1} = \kappa_{m2}$), and the unbroken edge mode emerges that disappears only at the zone boundary.

4.4 Edge modes in TI-2 regime

Let us now consider the TI-2 regime with $\Delta_B > 0$. Since the Hamiltonian $H(\mathbf{k})$ has only a linear term of Δ_B , the solution in the TI-2 regime can be obtained from that in the TI-1 regime by changing the sign of Δ_B . This conversion was indeed made in §3 for the straight edge.

In the zigzag edge, the spectrum given by eq.(63) has a negative slope because of $\Delta_B > 0$. Corresponding to eq.(39), the merging momentum $\kappa_m > 0$ is obtained from the condition

$$\lambda_- = -1, \quad (82)$$

instead of eq.(69). Then we obtain

$$\cos \kappa_m = \frac{\Delta_B \pm \sqrt{\Delta_B^2 - A^2(8 - A^2)}}{8 - A^2}, \quad (83)$$

instead of eq.(71), and

$$\Delta_{4c} = A\sqrt{8 - A^2} > 0, \quad (84)$$

instead of eq.(72). It is clear that the resultant solution of κ_m is the same as that in TI-1 range with the same $|\Delta_B|$. Similarly, one can check that all relevant quantities such as $\kappa_{m1}, \kappa_{m2}, \kappa_{c1}, \kappa_{c2}$ also have the same correspondence. The expression given by eqs.(77) and (80) remains valid, which means that the difference is now positive. This is naturally understood since the edge mode $E_\uparrow(\kappa)$ has the

negative slope.

5. Summary and discussion

We have analytically obtained spectrum and wave function of helical edge states for the BHZ model by identifying the annihilator for each case of (1,0) and (1,1) edges. The simplicity of the BHZ model has allowed us to obtain the complete information of the edge modes. Let us finally consider the case $A > 2$. In the straight edge, λ_{\pm} are always real in this case. Except for this difference, the property of the edge mode spectrum remains the same. In the zigzag edge, the case $A > \sqrt{8}$ allows no solution for merging momentum κ_m . This means the edge mode is present up to the zone boundary.

Edge spectrum shows different properties depending on edge geometry. In (1,0)-edge case, edge spectrum is proportional to $\sin k$. As the sign of Δ_B changes from negative to positive, the main location of edge mode moves from the center of the 1D BZ to the boundary. This movement is associated with the change of location of the bulk energy gap.

For the (1,1) edge, we have obtained the spectrum in the form of $\pm \Delta_B \sin \theta(\kappa)$, where $\theta(\kappa)$ is an odd function of momentum κ . Since $\theta(\kappa)$ does not depend on Δ_B as shown in eqs.(59) and (60), Δ_B appears only as the scale factor in the spectrum. With $\Delta_B = 0$, the edge modes become completely flat for the whole Brillouin zone, and two-fold degenerate as a consequence of the time-reversal symmetry.

The edge mode in the (1,1) geometry contains a novel reentrant part with extremely small binding energy. As seen from Fig.7, the binding energy for the reentrant part is only 10^{-3} of the overall energy. Except for many-body phenomena such as superconductivity and Kondo effect, we have been unaware of emergence of such extraordinary different energy scale. In spite of the tiny binding energy, the decay of the wave function toward inside the system looks quite normal, as judged by $|\lambda_{\pm}|$ which deviates clearly from unity.

Mathematically speaking, the tiny binding energy stems from the following factors:

- (i) The zone boundary $\kappa = \pi/2$ is a special point where the edge mode merges with bulk excitations.
- (ii) Group velocity of the edge mode is the same as that

of threshold excitation in the bulk at the merging point. Let us assume $\kappa \geq 0$ for simplicity. The condition (ii) requires the energy difference $E_{\uparrow} - E_b$ to be proportional to $(\kappa - \pi/2)^2$ near the zone boundary, and simultaneously to $(\kappa - \kappa_{m2})^2$ near the critical momentum. Hence in the intervening region $\kappa_{m2} < \kappa < \pi/2$, the growth of $E_{\uparrow} - E_b$ is constrained from both ends. Provided $\pi/2 - \kappa_{m2} \sim 0.1\pi$, we obtain the scaling factor $\sim 10^{-2}$ from the quadratic dependence of $E_{\uparrow} - E_b$. It is hoped that more physical explanation can be provided in the near future why the binding energy is so small.

- 1) C.L.Kane and E.J. Mele: Phys. Rev. Lett. 95, 146802 (2005).
- 2) C.L. Kane and E.J. Mele: Phys. Rev. Lett. 95, 226801 (2005).
- 3) L. Fu, C.L. Kane and E.J. Mele: Phys. Rev. Lett. 98, 106803 (2007).
- 4) L. Fu and C.L. Kane: Phys. Rev. B 76, 045302 (2007).
- 5) B.A. Bernevig and S.C. Zhang: Phys. Rev. Lett. 96, 106802 (2006).
- 6) B. A. Bernevig, T. L. Hughes and S.-C. Zhang: Science 314, 1757 (2006).
- 7) H. Zhang, C.-X. Liu, X.-L. Qi, X. Dai, Z. Fang and S.C. Zhang: Nat. Phys. 5, 438 (2009).
- 8) M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L.W. Molenkamp, X.-L. Qi, and S.C. Zhang: Science 318, 766 (2007).
- 9) D. Hsieh, D. Qian, L. Wray, Y. Xia, Y.S. Hor, R.J. Cava and M.Z. Hasan: Nature 452, 970 (2008).
- 10) Y. Xia, D. Qian, D. Hsieh, L. Wray, A. Pal, H. Lin, A. Bansil, D. Grauer, Y.S. Hor, R.J. Cava and M.Z. Hasan: Nat. Phys. 5, 398 (2009).
- 11) M. König, H. Buhmann, L.W. Molenkamp, T.L. Hughes, C.X. Liu, X.-L. Qi and S.C. Zhang: J. Phys. Soc. Jpn 77, 031007 (2008).
- 12) C. Wu, B.A. Bernevig, and S.C. Zhang: Phys. Rev. Lett. 96, 106401 (2006).
- 13) C. Xu and J. Moore, Phys. Rev. B 73, 045322 (2006).
- 14) B. Zhou, H.-Z. Lu, R.-L. Chu, S.-Q. Shen, and Q. Niu: Phys. Rev. Lett. 101, 246807 (2008).
- 15) E.B. Sonin: arXiv:1006.5218.
- 16) M. Fujita, K. Wakabayashi, K. Nakada, and K. Kusakabe: J. Phys. Soc. Jpn. 65, 1920 (1996).
- 17) K. Nakada and M. Fujita: Phys. Rev. B 54, 17954 (1996).
- 18) K. Imura, A. Yamakage, S. Mao, A. Hotta and Yoshio Kurokawa: arXiv:1004.5019; Phys. Rev. B in press.
- 19) M. Creutz and I. Horvath: Phys. Rev. D 50, 2297 (1994).
- 20) M. Creutz: Rev. Mod. Phys. 73, 119 (2001).
- 21) H.B. Nielsen and M. Ninomiya: Nucl. Phys. B 185 20 (1981); ibid. 193 173 (1981).